Nonlinear Postprocessing of Model Output

Caren Marzban
CAPS, University of Oklahoma
and
Department of Statistics, University of Washington

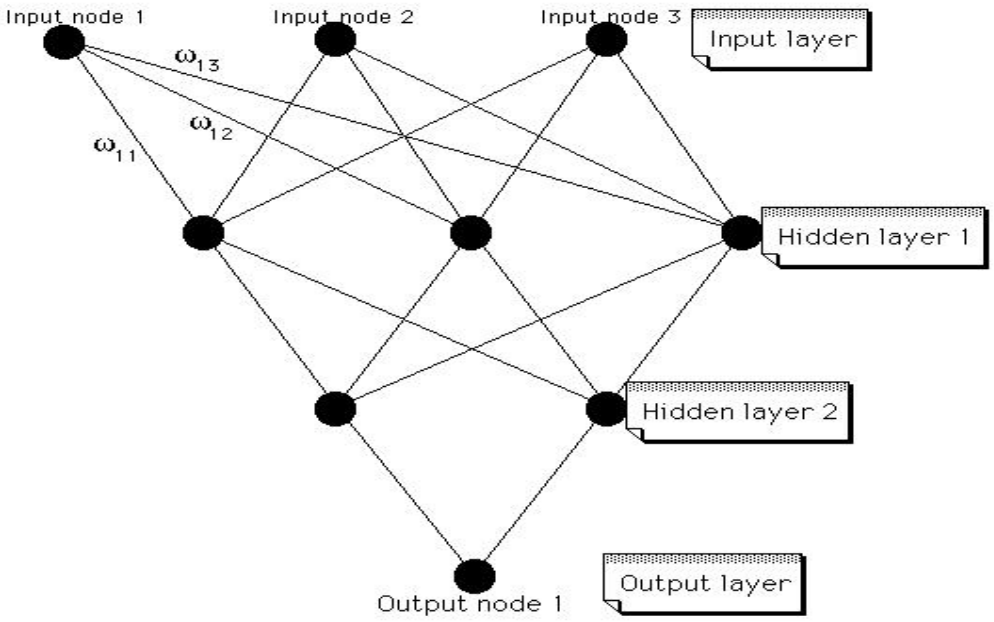
Outline

- Statistical Postprocessing ("Downscaling").
- Neural Networks + Bayes.
- ARPS, MRF $\longrightarrow$ temperature, ceiling/visibility.
- Heavy on Performance/Verification.

NN

$$y(x, \omega) = \omega x + \theta$$

$$y(x, \omega, H) = g\left(\sum_{i=1}^{H} \omega_i \; f(\sum_{j=1}^{N_{in}} \omega_{ij} x_j - \theta_j) - \omega\right)$$



Can learn any function if sufficient H.

H measures nonlinearity.

Things to worry about

Normalize distributions

Outliers

Normalize variables

Missing data

Filter out periodicity

Error function

input nodes

hidden nodes

Local minima

Overfitting

Error bars

Measure of performance:
MSE, MAE, bias, variance, scatterplots, residual plots, ROC plots, reliability plots, attributes diagrams, ...

Data

ARPS:

1 Time
2 Model forecast temp
3 Model relative humidity
4 Model wind direction
5 Model wind speed
6 Model sea-level pressure
7 Model cloud cover
8 Model precipitation rate
9 Model precipitation amount
10 OBSERVED TEMP

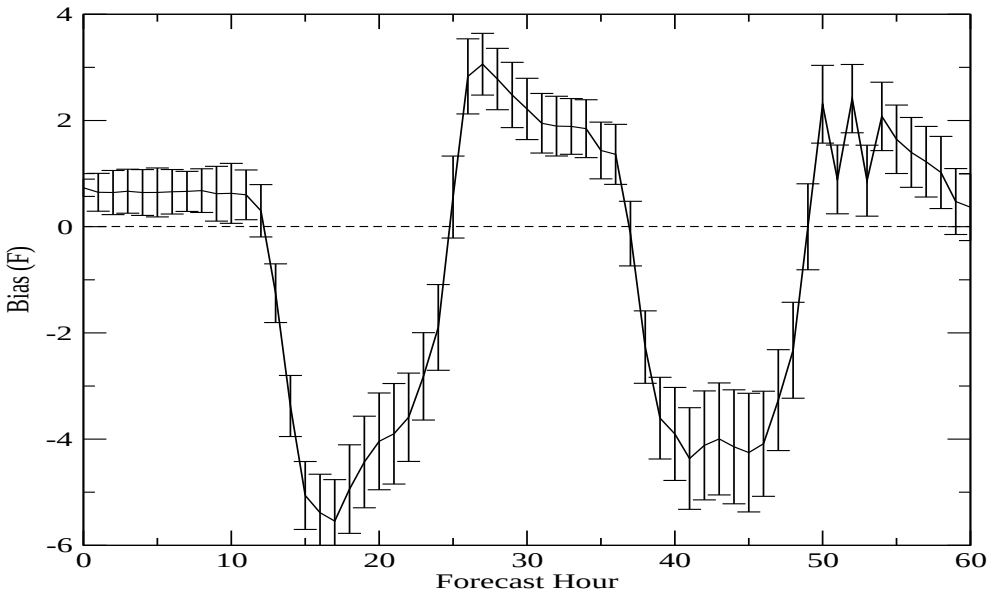
MRF:

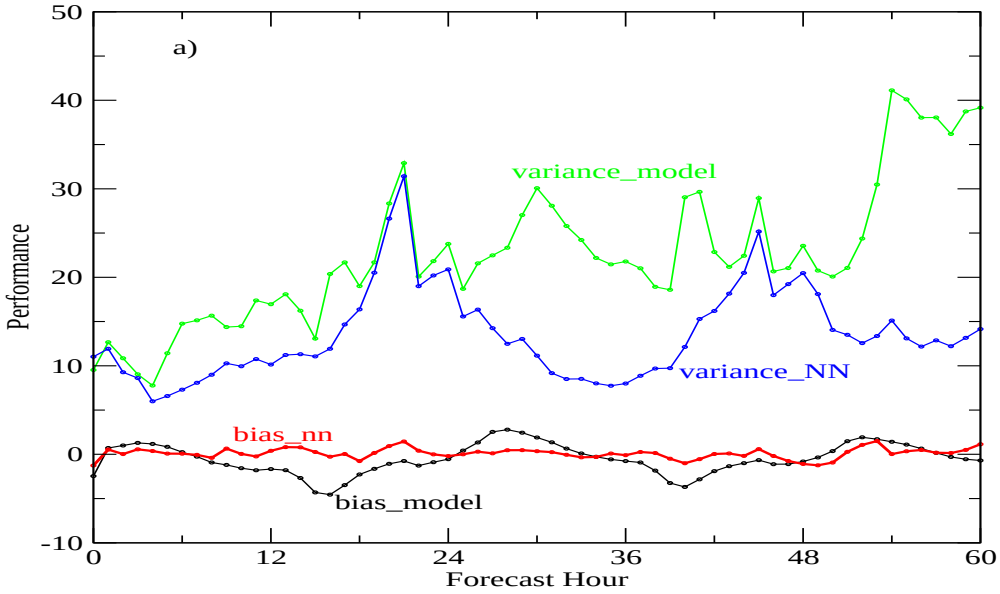
1 Time
2 Model forecast temp
3 Model relative humidity
4 Model precip amount in prev hour
5 Model cloud cover
6 OBSERVED TEMP

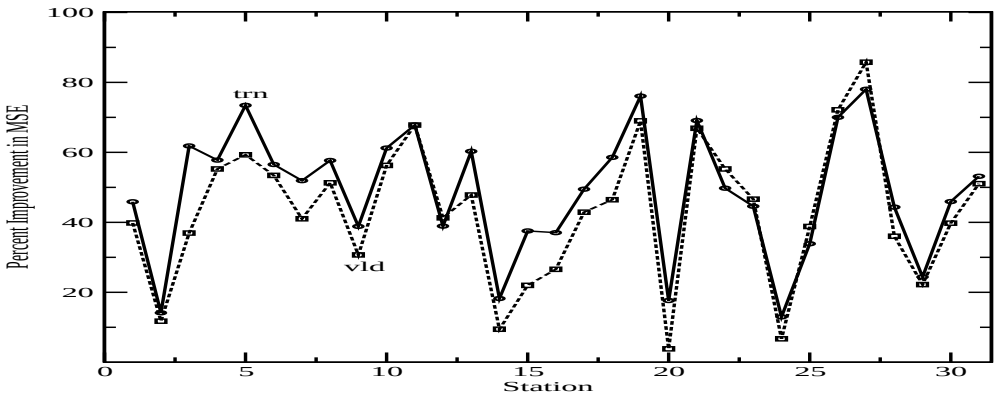
ARPS + Observation:

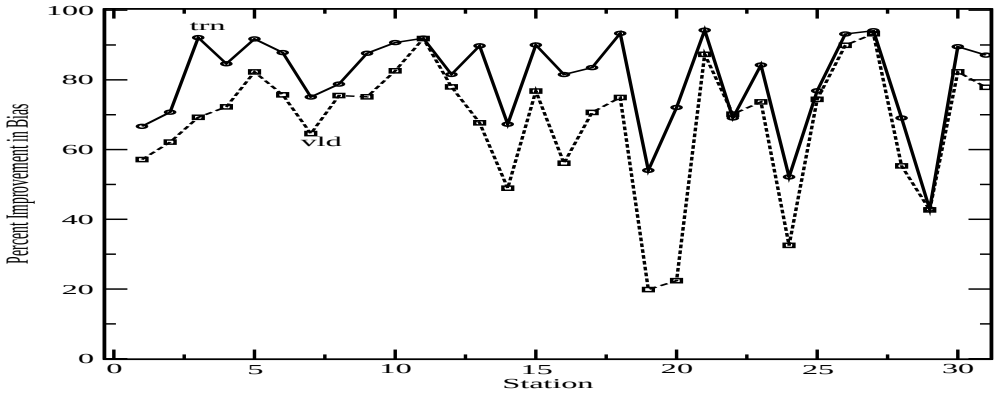
1 Time
2 Model forecast temp
3 Model relative humidity
4 Model wind direction
5 Model wind speed
6 Model sea-level pressure
7 Model cloud cover
8 Model precipitation rate
9 Model precipitation amount
10 OBSERVED TEMP
11 DEWPOINT
12 DEWPOINT DEPRESSION
13 HUMIDITY
14 CLOUD COVER
15 WIND_U
16 WIND_V
17 WIND GUST
18 MEAN SEA LEVEL PRESSURE
19 ALTIMETER
20 ALTIMETER
21 CLOUD AMOUNT
22 CLOUD HEIGHT
...
41 CEILING & VISIBILITY

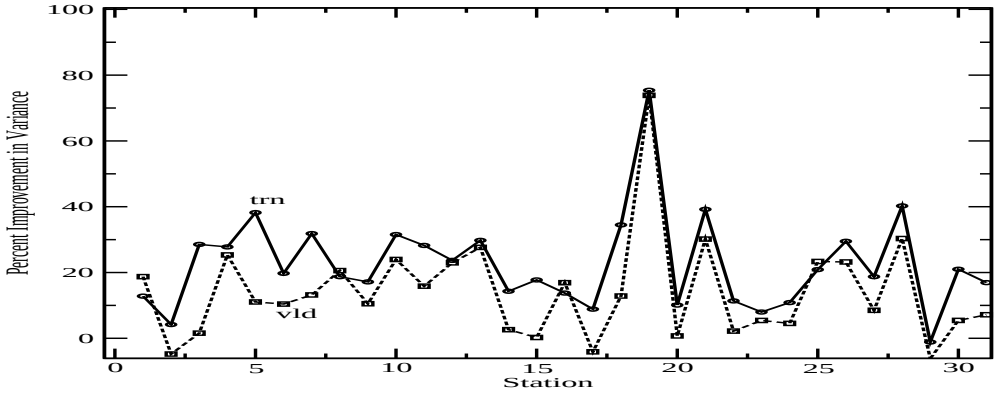
Verification

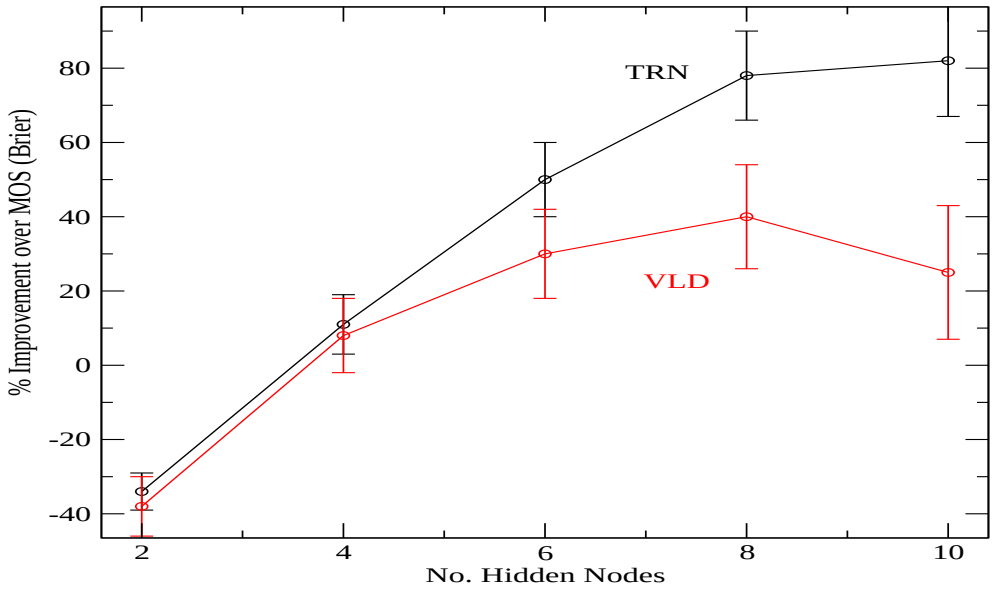


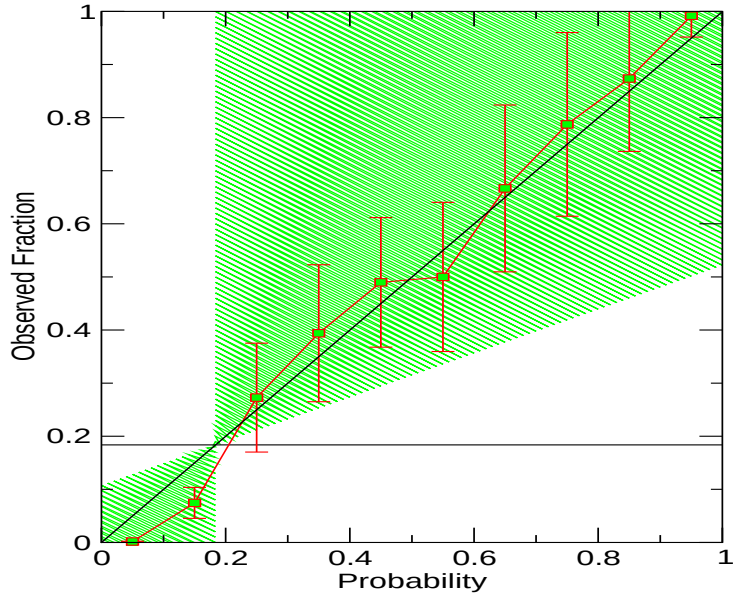



Percent Improvement in MSE
100
80
60
40
20
trn
vld
0
5
10
15
20
25
30
Station


Percent Improvement in Bias
100
80
60
40
20
0
trn
vld
0
5
10
15
20
25
30
Station


Percent Improvement in Variance
100
80
60
40
20
0
trn
vld
0
5
10
15
20
25
30
Station


% Improvement over MOS (Brier)
80
60
40
20
0
-20
-40
2
4
6
8
10
No. Hidden Nodes
TRN
VLD


Observed Fraction
1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1
Probability

Conclusion

NN Postprocessing:

- Improves NWP forecasts on "every" metric.
- Requires no supercomputers.
- 1-person operation.
- Quick development time.

Disadvantages?

- Little physical insight?
  (Not entirely true.)
  That's (nonlinear) life!

Next: MM5.